Appendix A. Supplementary material

Totally 11 pages including 10 Figure and 4 Tables.

Fig. S1 Paper-making wastewater treatment system water quality data.

Fig. S2 NLSTM hidden layer unit structure diagram.

Fig. S3 BNLSTM neural network structure diagram.

Fig. S4 The structure of the CNN-BNLSTM-Attention hybrid model.

Fig. S5 Dropout Strategy in Neural Net.

Fig. S6 CNN layer number experiment results for COD_{eff}.

Fig. S7 The comparison of actual value and prediction values for COD_{eff}.

Fig. S8 The comparison of actual value and prediction values for SS_{eff}.

Fig. S9 Back Attention Distribution Chart.

Fig. S10 Front Attention Distribution Chart.

Table S1. The hyperparameter search space settings

Table S2. Comparison of TPE, PSO search algorithms

Table.S3 Table of average attention weights for 50 runs of COD_{eff}

Table.S4 Table of average attention weights for 50 runs of SS_{eff}

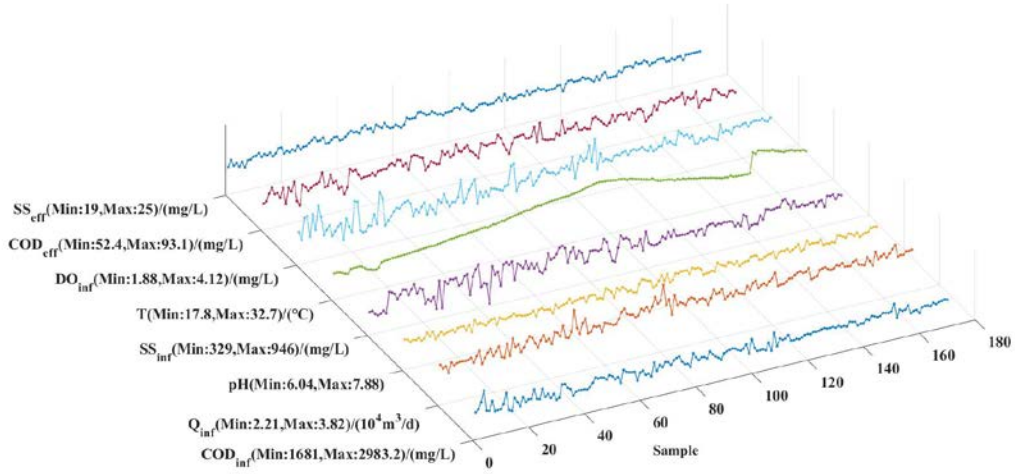

Fig. S1 Paper-making wastewater treatment system water quality data.

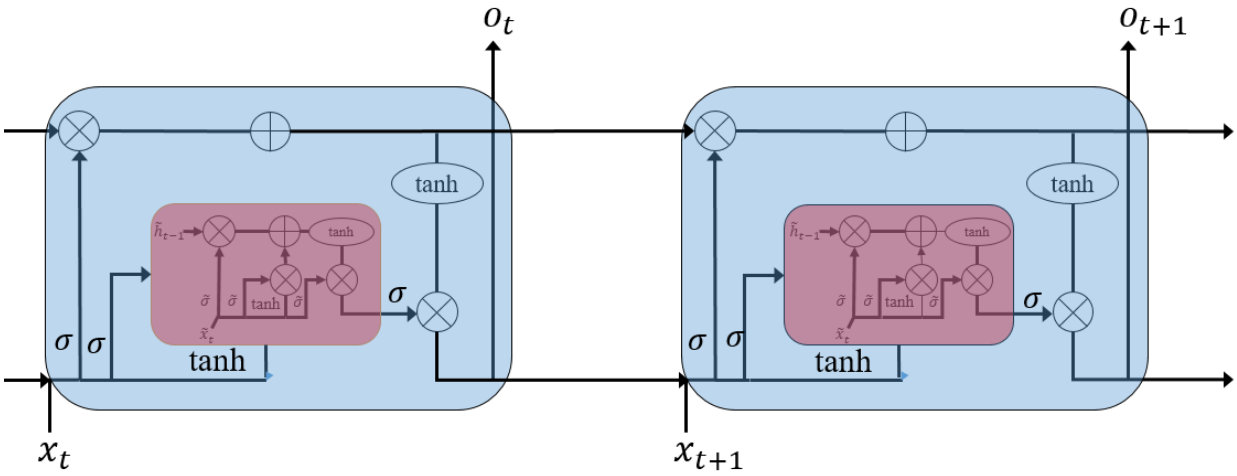

Fig. S2 NLSTM hidden layer unit structure diagram.

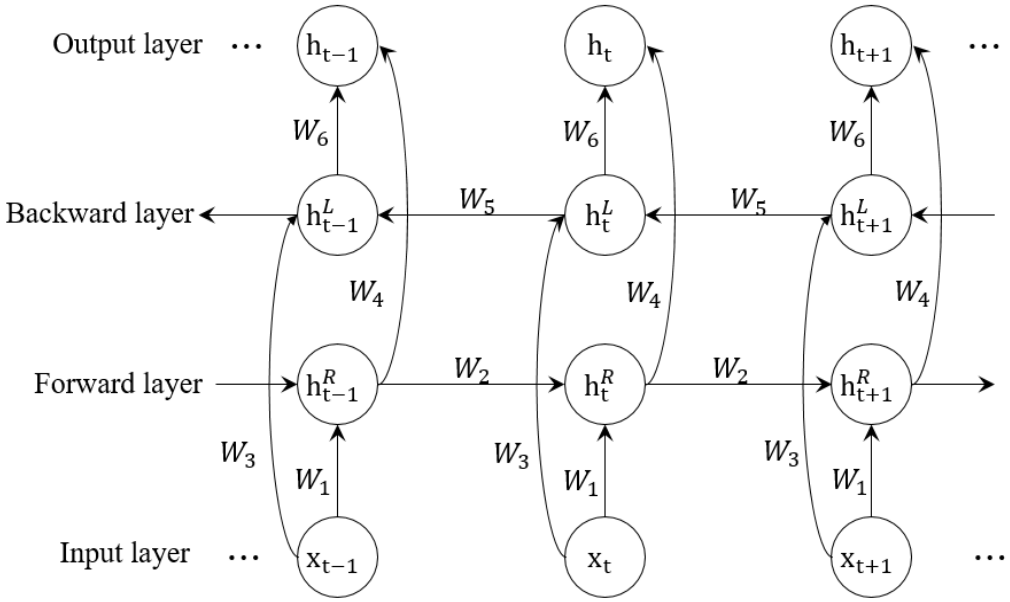

Fig. S3 BNLSTM neural network structure diagram.

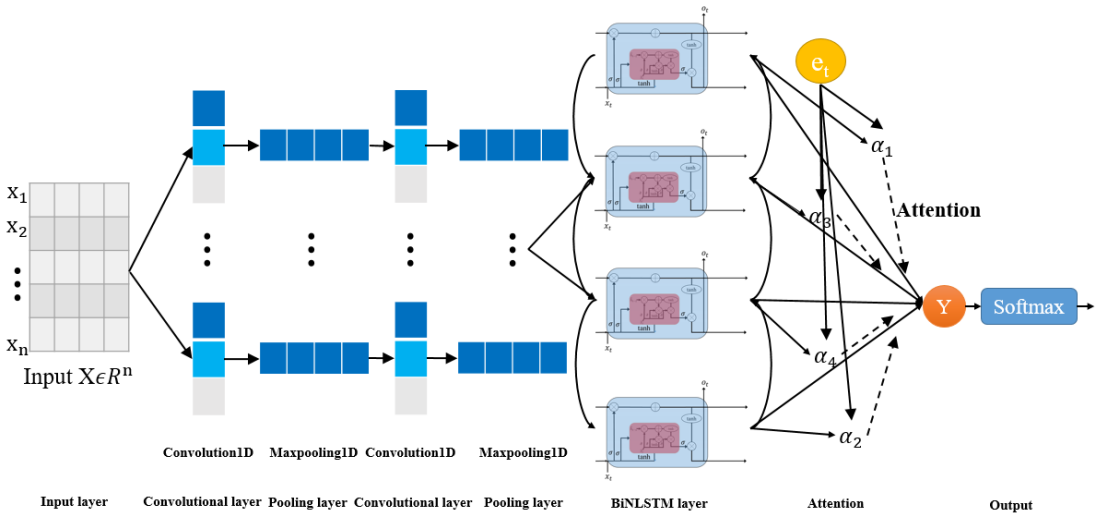

Fig. S4 The structure of the CNN-BNLSTM-Attention hybrid model.

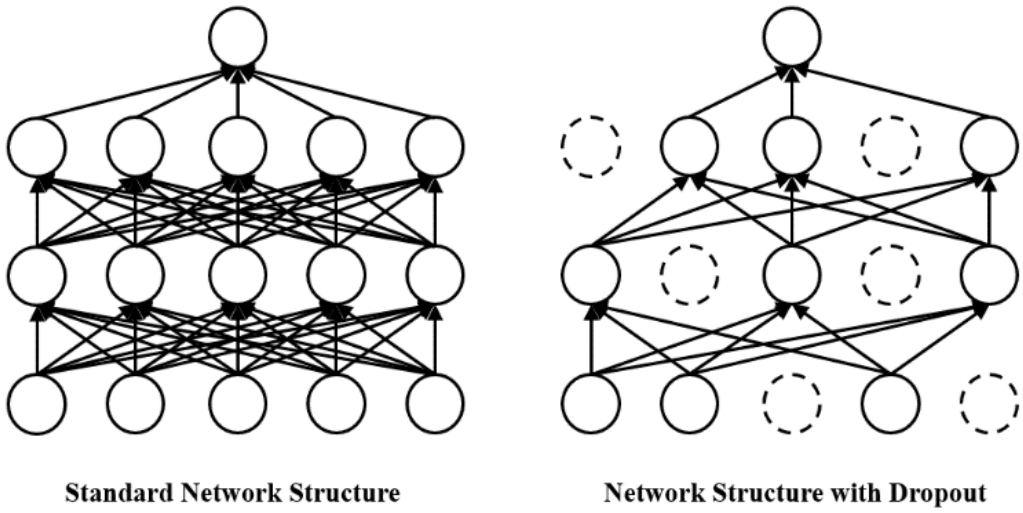

Fig. S5 Dropout Strategy in Neural Net.

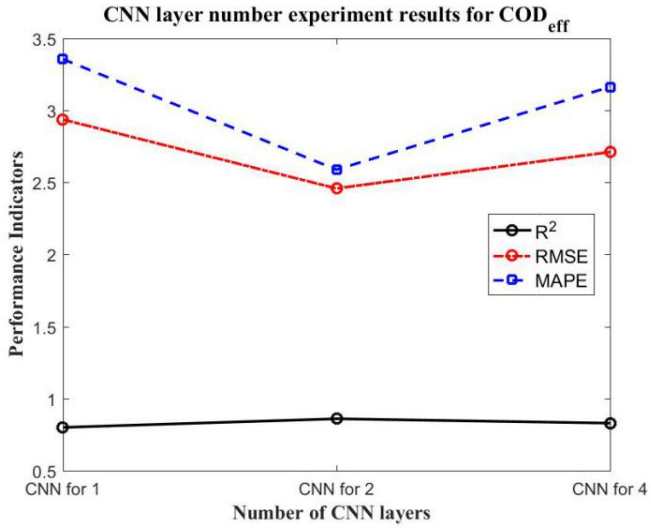

Fig. S6 CNN layer number experiment results for COD_{eff}.

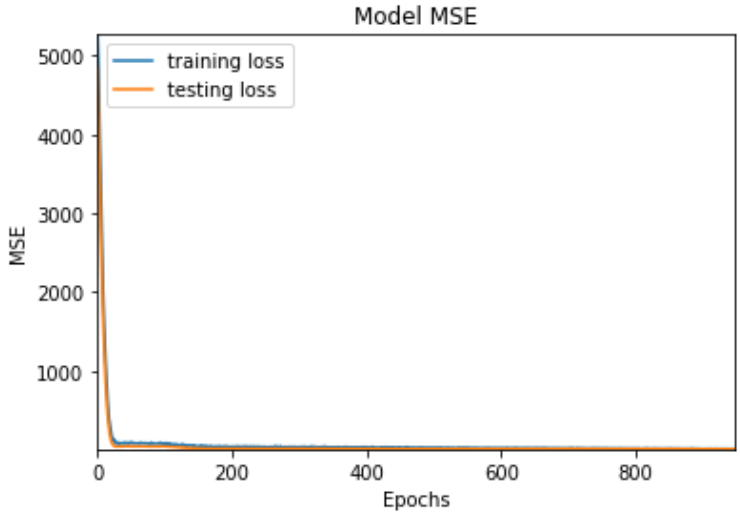

Fig. S7 The comparison of actual value and prediction values for COD_{eff}.

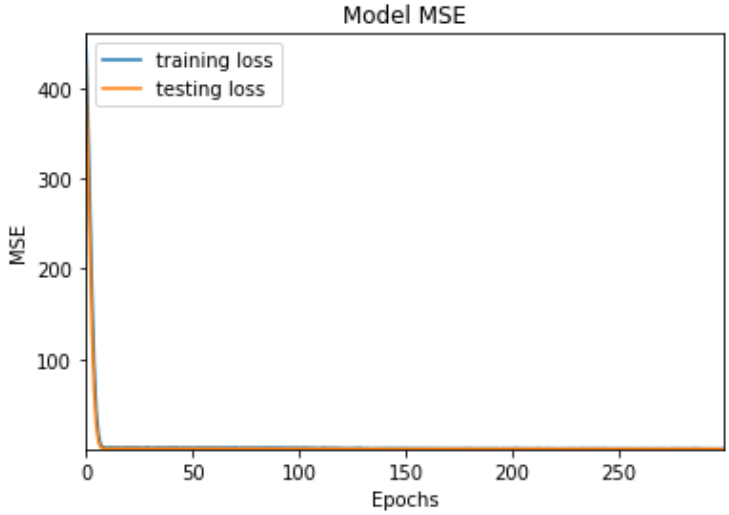

Fig. S8 The comparison of actual value and prediction values for SS_{eff}.

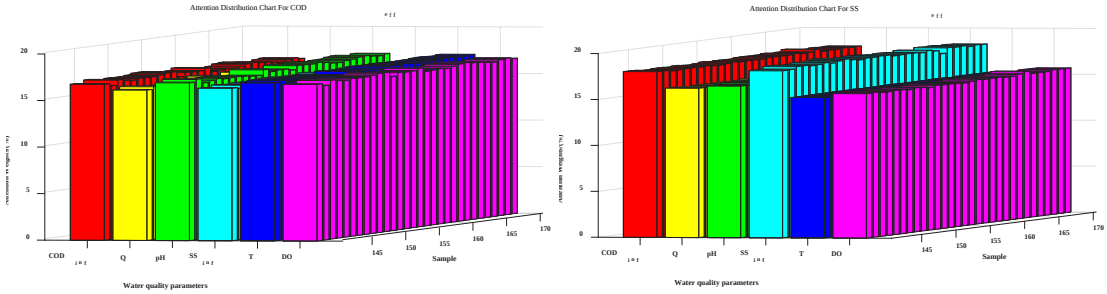

Fig. S9 Back Attention Distribution Chart.

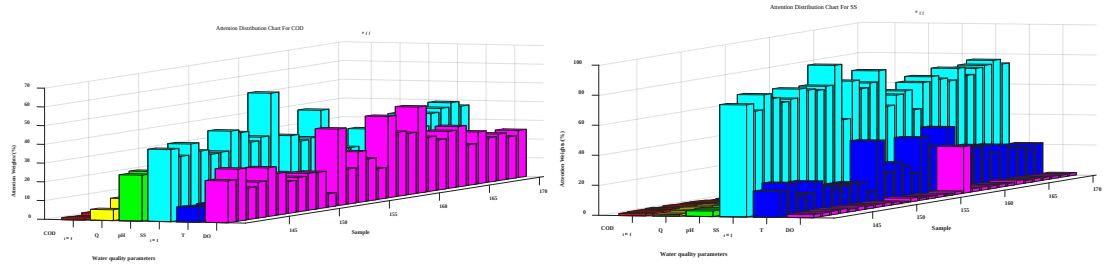

Fig. S10 Front Attention Distribution Chart.

Table.S1 The hyperparameter search space settings

Hyperparameter	Search scope
Filters-1	{32,48,64}
Filters-2	{32,48,64}
Kernel size	{1,3,5}
Number of hidden layer cells	[10,80]
Learning rate	[0.001,0.01]
Dropout	[0.005,0.5]
Batch size	[5,20]
Number of iterations	[700,1500]

Table.S2 Comparison of TPE, PSO search algorithms

Model	COD_{eff}				SS_{eff}			
	RMSE	MAPE	R^2	Time/min	RMSE	MAPE	R^2	Tim/min
TPE-CBNLSTMA	2.4586	2.5895	0.8595	21	0.4425	1.6037	0.7631	24
PSO-CBNLSTMA	3.2393	3.7266	0.7562	62	0.6014	2.0293	0.5625	71

Table.S3 Table of average attention weights for 50 runs of COD_{eff}

Run epoch	Average weight for COD_{eff}/%					
	COD_{inf}	Q	pH	SS_{inf}	T	DO
1~10	15.7	14.2	15.8	15.4	20.2	18.8
11~20	16.4	15.1	15.4	14.7	18.8	19.6
21~30	17.1	14.8	14.9	17.1	16.9	19.2
31~40	15.1	14.4	15.5	16.4	20.1	18.5
41~50	16.8	14.5	14.9	17.0	17.7	19.0

Table.S4 Table of average attention weights for 50 runs of SS_{eff}

Run epoch	Average weight for SS_{eff}/%					
	COD_{inf}	Q	pH	SS_{inf}	T	DO
1~10	17.9	17.3	16.5	16.5	15.6	16.1
11~20	17.3	17.0	16.9	16.9	15.3	16.6
21~30	18.3	16.7	16.5	17.2	15.0	16.4
31~40	17.0	17.2	16.4	16.3	16.0	17.1
41~50	17.5	16.1	17.1	17.2	15.3	16.9